



Recursive Estimation in Cointegrated VAR-Models

Johansen, Søren; Hansen, Henrik

Publication date:
1993

Document version
Peer reviewed version

Citation for published version (APA):
Johansen, S., & Hansen, H. (1993). *Recursive Estimation in Cointegrated VAR-Models*. (pp. 20). København.

DISCUSSION PAPERS

92-13

**Recursive Estimation
in Cointegrated VAR-Models**

Henrik Hansen & Søren Johansen

Institute of Economics, University of Copenhagen

Studiestræde 6, DK-1455 Copenhagen K, Denmark

tel. +45 33 91 21 66 - fax. +45 33 12 00 01

RECURSIVE ESTIMATION IN COINTEGRATED VAR-MODELS

by

Henrik Hansen^{*} & Søren Johansen^{**}

October 1992

^{*} Institute of Economics, University of Copenhagen

^{**} Institute of Mathematical Statistics, University of Copenhagen

Abstract

Some methods for the evaluation of parameter constancy in cointegrated VAR-models are discussed. Two different representations of the VAR-model are suggested; one in which all parameters in the model are estimated recursively, and another in which the short-run parameters are considered fixed and only the long-run parameters are estimated recursively. We suggest a procedure to evaluate the constancy of the estimated cointegration rank, and give a test of the constancy of the long-run parameters in the model for a given cointegration rank. Finally, the asymptotic distribution of the non-zero eigenvalues is given, and the time paths of these eigenvalues are graphed with pointwise asymptotic confidence bounds.

1. Introduction.

In the last decade there has been a growing interest in non-stationary time series; especially in the concepts of integrated and cointegrated time series. The basic aspects of cointegration are explained in the paper by Engle and Granger (1987) and the analysis of cointegration in the framework of vector autoregressive models have been treated by Reinsel and Ahn (1990), Johansen (1988,1991) and Johansen and Juselius (1990) among many others. The specific purpose of this paper is to consider graphical tests for parameter constancy in the cointegrated VAR-model by means of recursive estimation.

Recursive estimation is a widely used tool in the evaluation of parameter constancy. In general the procedure is applied in three different ways; forward recursions, backward recursions and windows of fixed length. The procedure used in this paper is the most common procedure which is forward recursions. In the forward recursion procedure the parameters of the model are estimated based on a subsample covering $t=1, \dots, T_0$ and the recursive formulae are used to update the parameter values stepwise from T_0 to the full sample values. The outcome of the recursive estimation is a sample of parameter estimates and often some summary statistics as well. The time paths of the estimated parameters and the summary statistics are presented graphically and used as diagnostic tools in the model evaluation. It is important to note that the null-hypothesis is parameter constancy and that we do not formulate a specific alternative. We regard the recursive analysis as a misspecification test where the purpose is to detect possible non-constancies in the parameters when there is no prior knowledge of structural breaks or time dependencies in the parameters.

In connection with the analysis of the cointegrated model we have two suggestions. Firstly, we suggest to treat the so called short-run parameters as fixed in the recursive estimation and secondly, we suggest to evaluate the time paths of the non-zero eigenvalues instead of all parameters in the model. Both suggestions are viewed as a mean to overcome the problems involved in the evaluation of the large number of estimated parameters in VAR-models.

Analyzing the model with fixed short-run dynamics is in contrast to the normal procedure when error-correction models are tested for parameter constancy. In most papers the long-run parameters are regarded as given, due to the superconsistency, and

the tests deals with the constancy of the short-run dynamics and the adjustment parameters¹. The main idea in this paper is to reverse this strategy and examine the long-run parameters, i.e. the cointegration relations under the assumption of constant short-run parameters.

The proposed test statistics are exemplified by a data set for the Danish wages analyzed in Hansen (1991). The data includes 8 endogenous series and the estimation period is 1973q1 to 1988q4. In the recursive analysis the initial period is 1973q1 to 1984q1. The empirical analysis in the present paper is only given as an illustration of the techniques.

The remainder of the paper is organized as follows. Section 2 gives a brief summary of the estimation technique and a discussion of the reason for fixing the short-run parameters. The constancy of the estimated cointegration rank is evaluated in section 3. In section 4 we propose an approximate test for the constancy of the cointegration space. Section 5 reports the asymptotic variance of the estimated non-zero eigenvalues and we discuss the information in the recursively estimated eigenvalues. Finally, section 6 contains some concluding remarks.

2. The cointegration model.

In this section the estimation method is briefly described, mainly to highlight the treatment of the short-run dynamics. The model considered is a p -dimensional, k 'th order VAR-model, written in error-correction form

$$\Delta X_t = \alpha \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Psi D_t + \mu + \epsilon_t, \quad (t=1, \dots, T). \quad (2.1)$$

The errors are assumed to be independent and Gaussian with mean zero and covariance matrix Ω , and the initial values X_{-k+1}, \dots, X_0 are fixed. D_t consist of n deterministic series and predetermined stationary, ergodic variables. The parameters are α and β ($p \times r$), $\Gamma_1, \dots, \Gamma_{k-1}$ ($p \times p$), Ψ ($p \times n$), μ ($p \times 1$) and Ω ($p \times p$) for some $r = 1, \dots, p$.

1. For examples of this procedure, see among others the special issue of *Journal of Policy Modeling* Vol. 14, 1992.

To ease the presentation we introduce some notation. Let $Z_{0t} = \Delta X_t$, $Z_{1t} = X_{t-1}$, $Z_{2t} = (\Delta X_{t-1}', \dots, \Delta X_{t-k+1}', D', 1)'$ and stack the parameters $(\Gamma_1, \dots, \Gamma_{k-1}, \Psi)$ in Γ . Using this notation the model is formulated as

$$Z_{0t} = \alpha \beta' Z_{1t} + \Gamma Z_{2t} + \epsilon_t \quad t = 1, \dots, T. \quad (2.2)$$

Maximum likelihood estimation of this model consists of a reduced rank regression² of Z_{0t} on Z_{1t} corrected for Z_{2t} . Thus, we define the residuals R_{0t} and R_{1t} by regression of ΔX_t and X_{t-1} on Z_{2t}

$$R_{0t} = Z_{0t} - M_{02} M_{22}^{-1} Z_{2t} \quad (2.3)$$

$$R_{1t} = Z_{1t} - M_{12} M_{22}^{-1} Z_{2t}$$

where

$$M_{ij} = T^{-1} \sum_{t=1}^T Z_{it} Z_{jt}' . \quad (2.4)$$

The remaining analysis can be based on the regression equation

$$R_{0t} = \alpha \beta' R_{1t} + \tilde{\epsilon}_t \quad t = 1, \dots, T, \quad (2.5)$$

where the product moment matrices are given by

$$S_{ij} = T^{-1} \sum_{t=1}^T R_{it} R_{jt}' . \quad (2.6)$$

The eigenvalues $1 > \hat{\lambda}_1 > \dots > \hat{\lambda}_p > 0$ are determined as solutions to the equation

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0, \quad (2.7)$$

and the eigenvectors $\hat{V} = (\hat{v}_1, \dots, \hat{v}_p)$ are normalized by $\hat{V}' S_{11} \hat{V} = I$.

The maximum likelihood estimators of β and α are given by

2. See Anderson (1951).

$$\hat{\beta} = (\hat{v}_1, \dots, \hat{v}_r), \quad \hat{\alpha} = S_{01} \hat{\beta}, \quad (2.8)$$

where it should be emphasized that only the space spanned by the vectors in β is estimable without further identifying restrictions in β .

A more detailed description of the estimation technique is given in Johansen (1988,1991) and Johansen & Juselius (1990), and the main reason for this short summary is to discuss the difference between the representation (2.2), which we will call the "Z-representation", and the representation (2.5), called the "R-representation". Although the formulation of the two representations is merely an analytical tool used to concentrate the likelihood function, the difference between the two representations is very useful in the recursive estimation because the estimation can be based on either of the two. If the recursive estimation is based on the original series, the Z-representation, all parameters are free to vary over time; if instead the residual series R_{0t} and R_{1t} are calculated using the full sample estimates of Γ , and the recursive estimation is performed using these series, the constancy of the parameters α , β is analyzed given the assumption of constant short-run dynamics.

In other words, it seems as if there are three "natural" approaches to the recursive analysis of parameter constancy in a cointegrated model. The first is to re-estimate all parameters as in the Z-representation, the second is to fix the short-run and re-estimate the long-run, the R-representation and finally the third approach is to fix the long-run parameters and re-estimate the short-run. The first and second approach is treated in this paper, while the third is done in many other papers, cf. footnote 1.

Another motivation for fixing the short-run parameters is that the unrestricted maximum likelihood procedure necessitates the inclusion of all relevant series in differences as well as levels. This often results in models with many insignificant parameters in Γ . Fixing the short-run parameters at the full sample values is one way of decreasing the variance in the estimated parameters and thereby centering the scope of the analysis to the long-run parameters.

3. The rank test.

The central parameter in the cointegration analysis is the cointegration rank, r , and although a structural break in the number of cointegrating relations is very unlikely, it

is important to have an idea of the degree of sample dependence in the estimated cointegration rank.

The cointegration rank is determined by using a likelihood ratio test, called the trace-test, given by

$$-2\ln(Q(r|p)) = -T \sum_{i=r+1}^p \ln(1 - \hat{\lambda}_i) \quad (3.1)$$

and the main result regarding this test is that the statistic converge in distribution to a multivariate version of the Dickey-Fuller distribution, as shown in Johansen (1988, 1991) and Reinsel & Ahn (1990).

In the recursive analysis another interesting result is that the $p-r$ smallest eigenvalues in (2.7) converge to zero at the rate of T^{-1} , while the r largest eigenvalues in (2.7) converge to the solution to

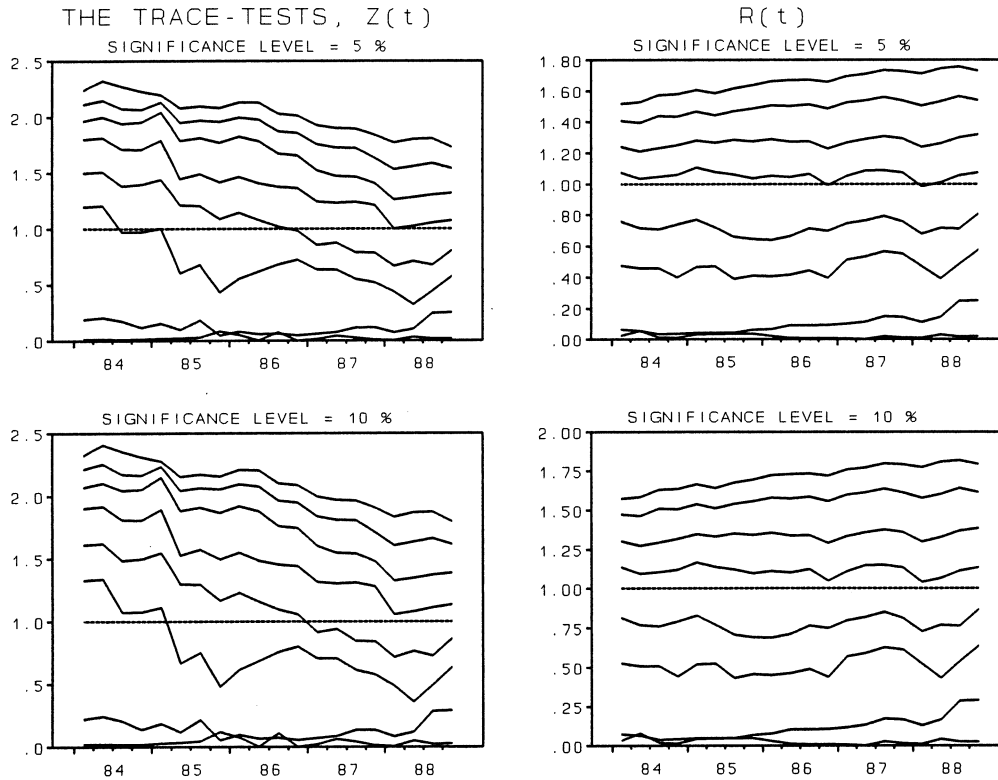
$$|\lambda \beta' \Sigma_{11} \beta - \beta' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta| = 0 \quad (3.2)$$

in which $\beta' \Sigma_{11} \beta$ is the asymptotic variance of $\beta' R_{1t}$, Σ_{00} is the asymptotic variance of R_{0t} and $\beta' \Sigma_{10}$ is the asymptotic covariancematrix for $\beta' R_{1t}$ and R_{0t} ³. This result implies that the rank test as a function of time will be upward sloping for $\hat{r} < r$ with a slope approximately equal to $\sum_{i=\hat{r}+1}^r \hat{\lambda}_i$, whereas the statistics are approximately constant for $\hat{r} \geq r$. Plotting the test statistics against time is therefore an auxiliary tool in the evaluation of the cointegration rank. It should be emphasized that graphical evaluations have to be done with care, but this plot is probably a good supplement to the formal tests in small and moderate samples.

In Figure 1 we show a plot of the trace-statistics against time. The left-hand side of the figure shows the statistics in the Z-representation, and the right-hand side shows the statistics based on the R-representation. All test statistics are scaled by a critical value, which means that values greater than unity imply rejection of the null-hypothesis, whereas values smaller than unity imply acceptance of the null. In the upper part of the figure the significance level is 5% and in the lower part it is 10%.

3. See appendix B.

Figure 1



The figure shows a remarkable difference between the two representations of the model. The test statistics calculated for the Z-representation are generally downward sloping, and the estimated cointegration rank takes values from 6 to 4 indicating a substantial sample dependence. In contrast to this result the statistics calculated from the R-representation have time paths in reasonable accordance with the theory, and the cointegration rank is constantly equal to 4, both at the 5% significance level and at the 10% significance level.

A problem is how to interpret these results. Our answer being that we are actually asking two different questions. In the Z-representation the question is: Which rank would have been estimated if we only had observations from 1 to t , where $t = T_0, \dots, T$, while the R-representation poses the question of the constancy of the cointegration rank given the full sample estimates of the short-run dynamics. In our view the latter question is the relevant one in the recursive analysis, which is why we conclude that the estimated cointegration rank is constant in this model.

4. A test for the constancy of the cointegration space.

Once the cointegration rank is determined, the model is estimated given the cointegration restriction, and this implies that the error-correction model can be evaluated applying most of the standard techniques, such as prediction tests. But tests based on the residuals or the recursive residuals give an evaluation of all parameters in the model, and even in the R-representation (2.5) we have not been able to isolate the long-run parameters, β , completely. Therefore, we suggest to use an approximate test for the constancy of the cointegration space. The approximate test is based on the likelihood ratio test for a known cointegration vector which is developed in Johansen & Juselius (1992).

We consider the hypothesis

$$H_\beta: sp(b) = sp(\beta), \quad (4.1)$$

where b is a known $p \times r$ matrix. In the recursive analysis we can perform a sequence of likelihood ratio tests of this hypothesis in which the estimate of β is based on different samples $1, \dots, t$, for $t = T_0, \dots, T$. Thus, we find the roots in

$$|\rho b' S_{11}(t)b - b' S_{10}(t) S_{00}^{-1}(t) S_{01}(t)b| = 0, \quad t = T_0, \dots, T, \quad (4.2)$$

and get a sequence of likelihood ratio statistics

$$-2\ln(Q(H_\beta | \hat{\beta}(t))) = t \sum_{i=1}^r \ln \left[\frac{1 - \hat{\rho}_i(t)}{1 - \hat{\lambda}_i(t)} \right], \quad t = T_0, \dots, T, \quad (4.3)$$

where $\hat{\rho}_i(t)$ are the roots of (4.2) and $\hat{\lambda}_i(t)$ are the roots of the unrestricted problem (2.7) with the moment matrices based on the sample $1, \dots, t$.

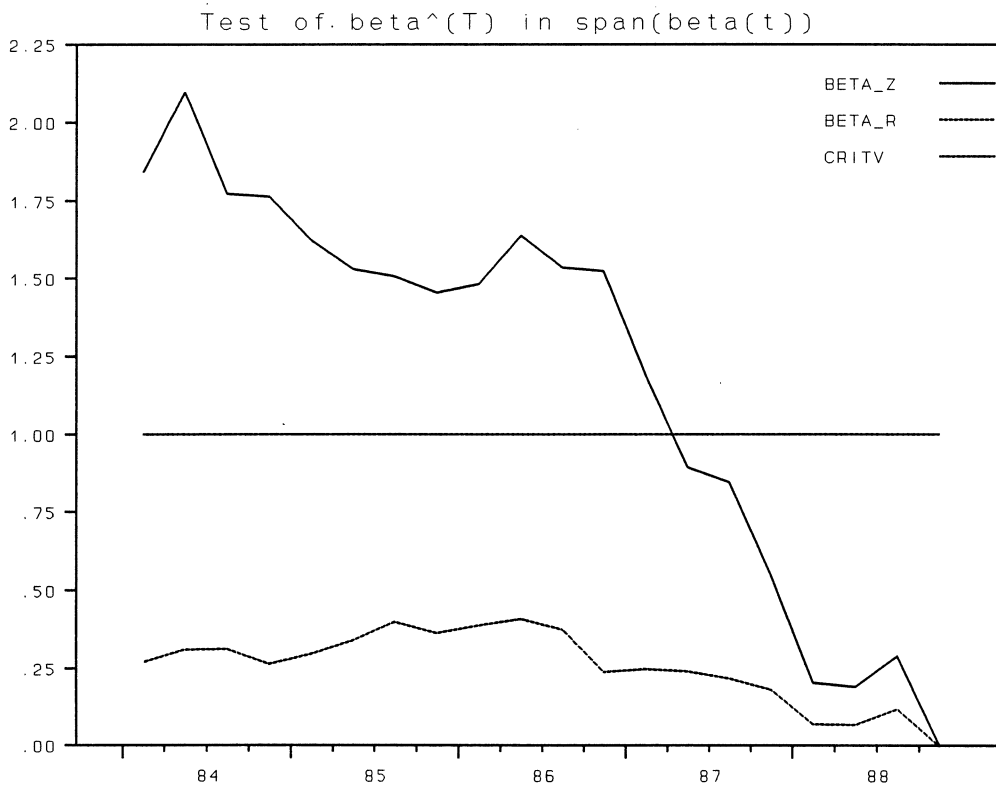
For each estimation period the LR-test has the same form, and in Johansen & Juselius (1992) it is shown that the statistic is asymptotically χ^2 distributed with $(p-r)r$ degrees of freedom. The test given in (4.2) has a much simpler form than the test given in Johansen & Juselius (1992), this is due to the fact that we claim to know the whole cointegration space instead of just a vector in the space.

The use of this sequence of tests as a test for parameter constancy can be motivated by the following. If a $p \times r$ matrix b is accepted to be in the space spanned by β , at some significance level α , this means that b is in the $(1-\alpha)$ 100% confidence region for

$\hat{\beta}$. If this holds for all estimated sub-samples we can not reject the hypothesis that β is constant. The remaining question is how to choose b , and we suggest to set $b = \hat{\beta}(T)$, where $\hat{\beta}(T)$ is the full-sample estimate of β . This choice is based on the fact that $\hat{\beta}(T)$ is the estimate with the smallest sample variation.

Figure 2 shows the approximate test calculated for both the Z- and the R-representation. In the figure the test statistics are scaled by the 5% critical value.

Figure 2



Once again we note a difference between the two representations. The hypothesis that $\hat{\beta}(T)$ is in the space spanned by $\beta(t)$ is rejected in the Z-representation, but accepted for all subperiods in the R-representation.

5. The time path of the eigenvalues.

In recursive estimations the summary statistics, such as (4.3), are often accompanied by plots of the time paths of the estimated parameters. In the cointegrated model it is unclear which parameters to plot because it is only the cointegration space that is

estimated. If the cointegration rank is greater than one the individual elements of β can only be identified if we impose identifying restrictions, therefore, we cannot investigate the constancy of the elements of β at this stage of the analysis. One possibility is to plot the elements of the estimated total impact matrix $\hat{\Pi} = \hat{\alpha}\hat{\beta}'$ since the elements of this matrix are identified. But the $\hat{\Pi}$ -matrix has dimension $p \times p$ giving rise to p^2 plots of parameter estimates. In the model analyzed in the present paper we would have to evaluate 64 plots of estimated parameters. Instead of plotting the elements of $\hat{\Pi}$ we suggest to plot the time paths of the r largest eigenvalues. If the estimated eigenvectors are normalized such that $\hat{\beta}'S_{11}\hat{\beta} = I$ we have a simple relation between the eigenvalues, the loadings and the cointegration vectors

$$\hat{\beta}'S_{10}S_{00}^{-1}S_{01}\hat{\beta} = \hat{\alpha}'S_{00}^{-1}\hat{\alpha} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_r) . \quad (5.1)$$

Thus, an evaluation of the time path of $\hat{\lambda}_i$ ($i=1, \dots, r$) can be seen as an evaluation of the i 'th column of $\hat{\alpha}$ or the i 'th column of $\hat{\beta}$, and structural changes in α or β will therefore be reflected in the estimated eigenvalues.

In order to evaluate the constancy of the eigenvalues we have to know the distribution of the estimated eigenvalues. In the following two Theorems we find the asymptotic distribution of the estimator for λ_1 under the assumption that λ_1 is a single root of the equation (2.7). The results can easily be modified to hold for any of the r largest distinct roots.

Theorem 1 determines the asymptotic distribution of $T^{1/2}(\hat{\lambda}_1 - \lambda_1)$, as being Gaussian with mean zero and variance matrix determined by $\text{Var}(S_{\epsilon\epsilon})$, $\text{Var}(S_{00})$ and $\text{Cov}(S_{\epsilon\epsilon}, S_{00})$. Theorem 2 gives an expression of the asymptotic variance.

THEOREM 1. *Under the assumption that λ_1 is a single root of eq. (2.7) it holds that the increment $\hat{\lambda}_1 - \lambda_1$ has the asymptotic expansion*

$$\hat{\lambda}_1 - \lambda_1 = -\alpha_1' \Sigma_{00}^{-1} (\lambda_1^{-1} (S_{\bar{\epsilon}\bar{\epsilon}} - \Omega) + (1 - \lambda_1^{-1}) (S_{00} - \Sigma_{00})) \Sigma_{00}^{-1} \alpha_1 + O_P(T^{-1}) . \quad (5.2)$$

The proof is given in appendix C.

THEOREM 2. Under the assumption that λ_1 is a single root of eq. (2.7) the asymptotic distribution of $T^{1/2}(\hat{\lambda}_1 - \lambda_1)$ is Gaussian with mean zero and variance given by

$$4(1 - \lambda_1)^2 \left(\sum_{m=1}^{\infty} \gamma_u(m)^2 - \sum_{m=1}^{\infty} \gamma_{uv}(m)^2 + \lambda_1 \right) \quad (5.3)$$

where γ_u and γ_{uv} are the autocovariance and crosscovariance function of the processes u_t and v_t given by

$$u_t = \lambda_1^{-1/2} \alpha_1' \Sigma_{00}^{-1} (\Delta X_t - E(\Delta X_t | Z_{2t})) \quad (5.4)$$

and

$$v_t = (\lambda_1(1 - \lambda_1))^{-1/2} \alpha_1' \Sigma_{00}^{-1} \epsilon_t \quad (5.5)$$

PROOF: It follows from Theorem 1 that we have the representation

$$\hat{\lambda}_1 - \lambda_1 = (1 - \lambda_1) T^{-1} \left(\sum_1^T \tilde{u}_t^2 - \sum_1^T \tilde{v}_t^2 \right) \quad (5.6)$$

where $\tilde{u}_t = \lambda_1^{-1/2} \alpha_1' \Sigma_{00}^{-1} R_{0t}$, and $\tilde{v}_t = (\lambda_1(1 - \lambda_1))^{-1/2} \alpha_1' \Sigma_{00}^{-1} \tilde{\epsilon}_t$.

The processes \tilde{u}_t and \tilde{v}_t converge weakly to u_t and v_t by the stationarity of Z_{0t} , Z_{2t} and ϵ_t , and we derive the result for the limiting processes u_t , v_t . That is, we investigate instead of (5.6)

$$\hat{\lambda}_1 - \lambda_1 = (1 - \lambda_1) T^{-1} \left(\sum_1^T u_t^2 - \sum_1^T v_t^2 \right) \quad (5.7)$$

We define $\gamma_u(m) = \text{Cov}(u_t, u_{t-m})$, $\gamma_v(m) = \text{Cov}(v_t, v_{t-m})$ and $\gamma_{uv}(m) = \text{Cov}(u_t, v_{t-m})$. In order to find the variance we note the following result which can be derived from Bartlett (1946) or Anderson (1971, Theorem 8.4.2., p.478)

$$T^{-1/2} \begin{bmatrix} \sum_t u_t^2 \\ \sum_t v_t^2 \end{bmatrix} \xrightarrow{w} N_2 \left(\begin{bmatrix} \gamma_u(0) \\ \gamma_v(0) \end{bmatrix}, 2 \sum_{m=-\infty}^{\infty} \begin{bmatrix} \gamma_u(m)^2 & \gamma_{uv}(m)^2 \\ \gamma_{vu}(m)^2 & \gamma_v(m)^2 \end{bmatrix} \right) \quad (5.8)$$

Since the process v_t is independent, identically distributed we find that the variance is $(\lambda_1(1-\lambda_1))^{-1} \alpha_1' \Sigma_{00}^{-1} \Omega \Sigma_{00}^{-1} \alpha_1 = 1$, see (A.6) and (A.8) in appendix A, and this implies that

$$\gamma_v(m) = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$$

Moreover $\gamma_{uv}(m) = 0$ if $m < 0$, and

$$\begin{aligned} \gamma_{uv}(0) &= \text{Var}(u_t, v_t) \\ &= \lambda_1^{-1} (1 - \lambda_1)^{-1/2} \alpha_1' \Sigma_{00}^{-1} E\{\epsilon_t (\Delta X_t - E(\Delta X_t | Z_{2t}))\}' \Sigma_{00}^{-1} \alpha_1 \\ &= \lambda_1^{-1} (1 - \lambda_1)^{-1/2} \alpha_1' \Sigma_{00}^{-1} \Omega \Sigma_{00}^{-1} \alpha_1 \\ &= (1 - \lambda_1)^{1/2} \end{aligned}$$

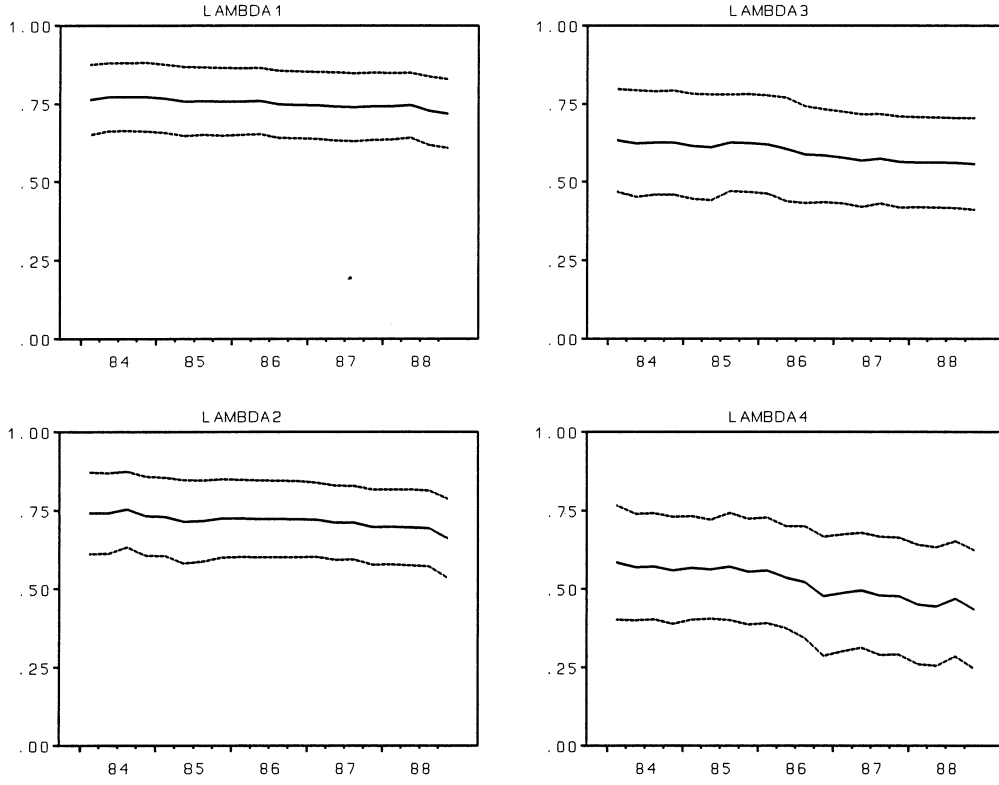
In terms of these variables the asymptotic variance of $T^{1/2}(\hat{\lambda}_1 - \lambda_1)$ is given by

$$\begin{aligned} &2(1 - \lambda_1)^2 \sum_{m=-\infty}^{\infty} [\gamma_u(m)^2 + \gamma_v(m)^2 - 2\gamma_{uv}(m)^2] = \\ &2(1 - \lambda_1)^2 \left[\left(1 + 2 \sum_{m=1}^{\infty} \gamma_u(m)^2 \right) + 1 - 2 \left(1 - \lambda_1 + \sum_{m=1}^{\infty} \gamma_{uv}(m)^2 \right) \right] = \\ &4(1 - \lambda_1)^2 \left[\lambda_1 + \sum_{m=1}^{\infty} \gamma_u(m)^2 - \sum_{m=1}^{\infty} \gamma_{uv}(m)^2 \right] \end{aligned}$$

□

In order to apply the results one has to estimate the asymptotic variance given by (5.3). The exact formula for the case of an autoregressive process looks forbidding, so the maximum likelihood estimator of the variance is difficult to calculate. Another possibility is to use a non-parametric estimator based upon the empirical autocovariance function of the process R_{0t} , and it seems a good idea to use a kernel estimator of the asymptotic variance.

Figure 3



An example is the Bartlett kernel which suggests to use

$$4(1-\hat{\lambda}_1)^2(\hat{\lambda}_1 + \sum_{h=1}^M (1-h/M)^2(r_u(h)^2 - r_{uv}(h)^2)) \quad (5.9)$$

where

$$\begin{aligned} r_u(h) &= T^{-1} \sum_{t=h}^T \hat{u}_t \hat{u}_{t-h} , \\ r_{uv}(h) &= T^{-1} \sum_{t=h}^T \hat{u}_t \hat{v}_{t-h} \end{aligned} \quad (5.10)$$

for

$$\hat{u}_t = \hat{\lambda}_1^{-1/2} \hat{\alpha}_1' S_{00}^{-1} R_{0t} \quad (5.11)$$

and

$$\hat{v}_t = (\hat{\lambda}_1(1 - \hat{\lambda}_1))^{-1/2} \hat{\alpha}_1' S_{00}^{-1} \hat{\epsilon}_t = (\hat{\lambda}_1(1 - \hat{\lambda}_1))^{-1/2} \hat{\alpha}_1' S_{00}^{-1} (R_{0t} - \hat{\alpha} \hat{\beta}' R_{1t}) \quad (5.12)$$

Figure 3 shows the time paths of the eigenvalues with 95%-confidence bounds. In the example we have used the Bartlett kernel where the bandwidth, M , has been set equal to 4, and the eigenvalues are estimated based on the R-representation. All the eigenvalues are constant, in particular there is no sign of any break-points or significant drift. The plot of the eigenvalues support the conclusion reached in section 4 where we accepted the hypothesis of a constant cointegration space.

6. Concluding remarks.

This paper addresses the issue of testing for the constancy of the long-run parameters in a cointegrated VAR-model. It is shown that the maximum likelihood estimation of cointegrated VAR-models with Gaussian errors leads to two "natural" representations of the model which can be used in the recursive analysis. In the first representation all parameters in the model are re-estimated in each period whereas only the long-run parameters are re-estimated in the second representation. The two representations shows substantially different results in the example used in the paper, and therefore the empirical purpose of the VAR-model becomes important. The requirement must be that all parameters are constant if the main purpose of the model is prediction of all variables in the system. On the other hand it seems a good idea to fix some of the parameters in order to narrow the scope to the central parameters if the VAR-model is the first step in a structural analysis and the model as such is over-parametrized.

The main issue in the paper is to present tests for the constancy of the cointegration space and we suggest two ways of testing. The two tests do not require an identification of the individual cointegration vectors and they are therefore applicable at an early stage of the empirical analysis.

Appendix A. Some technical results.

This appendix reports some technical results for easy reference. The results can be found in Johansen (1991 appendix A) and they are stated here without proofs. In Lemma A.1. we give some moment relations, which simplify the calculations when the

β 's are normalized such that $\beta' \Sigma_{11} \beta = \mathbf{I}$ and $\beta' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$.

Lemma A.2 and A.3 summarize the properties of the process X_t .

LEMMA A.1. *The product moment matrices satisfy the relations*

$$S_{01} \beta = \alpha \beta' S_{11} \beta + S_{\epsilon 1} \beta \quad (\text{A.1})$$

$$S_{00} = \alpha \beta' S_{11} \beta \alpha' + S_{\tilde{\epsilon} \tilde{\epsilon}} + S_{\epsilon 1} \beta \alpha' + \alpha \beta' S_{1\epsilon'} \quad (\text{A.2})$$

where

$$S_{\tilde{\epsilon} \tilde{\epsilon}} = T^{-1} \sum_{t=1}^T \tilde{\epsilon}_t \tilde{\epsilon}_t', \quad \text{and} \quad S_{\epsilon 1} = T^{-1} \sum_{t=1}^T \tilde{\epsilon}_t R_{1t}' = T^{-1} \sum_{t=1}^T \epsilon_t R_{1t}'.$$

When β is chosen such that $\beta' \Sigma_{11} \beta = \mathbf{I}$ and $\beta' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta = \Lambda$ then

$$\Sigma_{01} \beta = \alpha \beta' \Sigma_{11} \beta = \alpha \quad (\text{A.3})$$

$$\Sigma_{00} = \alpha \beta' \Sigma_{11} \beta \alpha' + \Omega = \alpha \alpha' + \Omega \quad (\text{A.4})$$

$$\Sigma_{00}^{-1} = \Omega^{-1} - \Omega^{-1} \alpha (\mathbf{I} + \alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \quad (\text{A.5})$$

$$\alpha' \Sigma_{00}^{-1} \alpha = \Lambda \quad (\text{A.6})$$

$$\alpha' \Omega^{-1} \alpha = \Lambda (\mathbf{I} - \Lambda)^{-1} \quad (\text{A.7})$$

$$\alpha' \Sigma_{00}^{-1} = (\mathbf{I} - \Lambda) \alpha' \Omega^{-1} \quad (\text{A.8})$$

$$\Sigma_{00}^{-1} - \Sigma_{00}^{-1} \alpha (\alpha' \Sigma_{00}^{-1} \alpha)^{-1} \alpha' \Sigma_{00}^{-1} = \alpha_{\perp} (\alpha'_{\perp} \Omega \alpha_{\perp})^{-1} \alpha'_{\perp} \quad (\text{A.9})$$

LEMMA A.2. Define $\tau = C\mu$ and let $\gamma(p \times (p-r-1))$ be chosen orthogonal to τ and β , such that (β, γ, τ) span all of \mathbb{R}^p . Then it follows from the moving average representation of the process X_t that

$$T^{-1/2}\gamma'X_{[Tu]} \rightarrow^w \gamma'CW(u), \quad \text{for } T \rightarrow \infty \text{ and } u \in [0,1] \quad (\text{A.10})$$

$$T^{-1}\tau'X_{[Tu]} \rightarrow^P \tau'\tau u, \quad \text{for } T \rightarrow \infty \text{ and } u \in [0,1] \quad (\text{A.11})$$

where $C = \beta_\perp (\alpha'_\perp (\mathbf{I} - \sum_1^{k-1} \Gamma_i) \beta_\perp)^{-1} \alpha'_\perp$, and $W(u)$ is the Brownian motion defined by

$$T^{-1/2} \sum_{i=0}^{[Tu]} \epsilon_i \rightarrow^w W(u).$$

LEMMA A.3. Define $B_T = (\bar{\gamma}, T^{-1/2}\bar{\tau})$, and $G' = (G'_1, G'_2)$ where

$$G_1(t) = \bar{\gamma}'C\left(W(t) - \int_0^1 W(u)du\right), \quad G_2(t) = t^{-1/2}, \quad (t \in [0,1]), \text{ then}$$

$$B_T'(S_{10} - S_{11}\beta\alpha') \rightarrow^w \int_0^1 G(dW)' \quad (\text{A.12})$$

$$T^{-1}B_T'S_{11}B_T \rightarrow^w \int_0^1 GG'du \quad (\text{A.13})$$

and

$$\beta'S_{11}\beta \rightarrow^P \beta'\Sigma_{11}\beta = \text{Var}(\beta'X_t|Z_{2t}) \quad (\text{A.14})$$

$$\beta'S_{10} \rightarrow^P \beta'\Sigma_{10} = \text{Cov}(\beta'X_t, \Delta X_t|Z_{2t}) \quad (\text{A.15})$$

$$S_{00} \rightarrow^P \Sigma_{00} = \text{Var}(\Delta X_t|Z_{2t}) \quad (\text{A.16})$$

Appendix B. The asymptotic properties of the equation (2.7).

The results given in this appendix have previously been stated in Johansen (1988, 1991).

We want to apply the asymptotic properties for the process X_t to discuss the asymptotic behavior of the equation

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0 \quad (\text{B.1})$$

and hence the roots.

We multiply (B.1) by the matrix $A_T = (\beta, T^{-1/2} B_T)$ and its transposed to get

$$\left| \lambda \begin{bmatrix} \beta' S_{11} \beta & T^{-1/2} \beta' S_{11} B_T \\ T^{-1/2} B_T' S_{11} \beta & T^{-1} B_T' S_{11} B_T \end{bmatrix} - \begin{bmatrix} \beta' S_{10} \\ T^{-1/2} B_T' S_{10} \end{bmatrix} S_{00}^{-1} \begin{bmatrix} \beta' S_{10} \\ T^{-1/2} B_T' S_{10} \end{bmatrix}' \right| = 0 \quad (\text{B.2})$$

For $T \rightarrow \infty$ we apply Lemma A.3 and Lemma A.1 to get the limit

$$\begin{aligned} & \left| \lambda \begin{bmatrix} \beta' \Sigma_{11} \beta & 0 \\ 0 & \int_0^1 G G' du \end{bmatrix} - \begin{bmatrix} \beta' \Sigma_{10} \\ 0 \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \beta' \Sigma_{10} \\ 0 \end{bmatrix}' \right| = \left| \begin{bmatrix} \lambda I - \Lambda & 0 \\ 0 & \lambda \int_0^1 G G' du \end{bmatrix} \right| \quad (\text{B.3}) \\ & = \prod_{i=1}^r (\lambda - \lambda_i) \lambda^{p-r} \left| \int_0^1 G G' du \right| = 0 \end{aligned}$$

This equation has $p-r$ zero roots, and the remaining roots are the roots of (3.2).

This shows that the $p-r$ smallest eigenvalues of (B.1) tend to zero and the r largest tend to the solution of (3.2).

Appendix C. Proof of the results in section 5.

In this appendix we first give a Lemma that contains an expansion of a determinant when the first term is zero. This Lemma is used in the proof of Theorem 1.

LEMMA C.1. *Let U and V be symmetric positive semidefinite matrices. If U has eigenvalues $\rho_1 > \dots > \rho_p = 0$, and eigenvectors v_1, v_2, \dots, v_p then*

$$f(t) = |U + tV| = t \left(\prod_{i=1}^{p-1} \rho_i \right) v_p' V v_p + O(t^2)$$

PROOF: Multiply the matrices by $V = (v_1, v_2, \dots, v_p)$ and let $v_{ij} = v_i' V v_j$. Then

$$f(t) = \left| \begin{bmatrix} (\rho_1 + tv_{11}) & tv_{12} & \cdots & tv_{1p} \\ tv_{21} & (\rho_2 + tv_{22}) & \cdots & tv_{2p} \\ \vdots & & \ddots & \\ tv_{p1} & tv_{p2} & \cdots & tv_{pp} \end{bmatrix} \right| = \rho_1 \cdots \rho_{p-1} tv_{pp} + O(t^2) \quad \square$$

Proof of Theorem 1.

In order to ease the notation in the proof we introduce the following:

Denote (B.2) and (B.3) by

$$|\lambda \hat{A} - \hat{B}| = 0 \quad (\text{C.1})$$

$$|\lambda A - B| = 0 \quad (\text{C.2})$$

Introduce the function

$$h(\lambda) = |\lambda \hat{A} - \hat{B}| \quad (\text{C.3})$$

and define the increments

$$\partial \lambda_1 = \hat{\lambda}_1 - \lambda_1, \quad \partial A = \hat{A} - A, \quad \partial B = \hat{B} - B. \quad (\text{C.4})$$

The eigenvalue λ_1 is a solution of $h(\hat{\lambda}) = 0$, and the proof consists of expanding this equation around the point λ_1 . Let us write (B.1) in the form

$$\begin{aligned} 0 &= h(\hat{\lambda}_1) = |\hat{\lambda}_1 \hat{A} - \hat{B}| \\ &= |(\lambda_1 + \partial \lambda_1)(A + \partial A) - (B + \partial B)| \\ &= |(\lambda_1 A - B) + (\partial \lambda_1 A + \lambda_1 \partial A - \partial B + \partial \lambda_1 \partial A)| \end{aligned}$$

The matrix $(\lambda_1 A - B)$ is singular with one eigenvalue equal to 0 and the corresponding eigenvector is the p -dimensional unit vector e_1 , since

$$\lambda_1 A e_1 - B e_1 = (\lambda_1 \beta' \Sigma_{11} \beta - \beta' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta) \tilde{e}_1 = (\lambda_1 I - \Lambda) \tilde{e}_1 = 0,$$

where \tilde{e}_1 is the r -dimensional unit vector.

We now apply Lemma C.1 and (B.3) to find the expansion of $h(\hat{\lambda}_1)$ in a neighborhood of λ_1 :

$$0 = h(\hat{\lambda}_1) = \prod_{i=2}^r (\lambda_1 - \lambda_i) \lambda_1^{p-r} \left| \int_0^1 G G' du \right| (\partial \lambda_1 + \lambda_1 e_1' \partial A e_1 - e_1' \partial B e_1) + O_P(T^{-1})$$

Thus the first order expansion of $\hat{\lambda}_1 - \lambda_1$ is given by

$$\begin{aligned} \partial \lambda_1 &= -\lambda_1 \partial(\beta_1' \Sigma_{11} \beta_1) + \partial(\beta_1' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta_1) + O_P(T^{-1}) \\ &= -\lambda_1 (\beta_1' S_{11} \beta_1 - 1) + (\beta_1' S_{10} S_{00}^{-1} S_{01} \beta_1 - \lambda_1) + O_P(T^{-1}) \end{aligned} \quad (C.5)$$

Next we derive an expression for $\hat{\lambda}_1 - \lambda_1$ which only involves the differential of Ω and Σ_{00} . We apply the results of Lemma A.1, and find from (A.1), (A.2), (A.3) and (A.4) that

$$\partial \Sigma_{01} \beta = \alpha \partial(\beta' \Sigma_{11} \beta) + S_{\epsilon 1} \beta \quad (C.6)$$

$$\partial \Sigma_{00} = \alpha \partial(\beta' \Sigma_{11} \beta) \alpha' + \partial \Omega + S_{\epsilon 1} \beta \alpha' + \alpha \beta' S_{1\epsilon} \quad (C.7)$$

From (C.7) we find that

$$\begin{aligned} \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{00}) \Sigma_{00}^{-1} \alpha_1 &= \\ \lambda_1^2 \partial(\beta_1' \Sigma_{11} \beta_1) + \alpha_1' \Sigma_{00}^{-1} \partial \Omega \Sigma_{00}^{-1} \alpha_1 + \lambda_1 (\alpha_1' \Sigma_{00}^{-1} S_{\epsilon 1} \beta_1 + \beta_1' S_{1\epsilon} \Sigma_{00}^{-1} \alpha_1) \end{aligned} \quad (C.8)$$

and

$$\partial \beta_1' \Sigma_{10} \Sigma_{00}^{-1} \Sigma_{01} \beta_1 = (\partial \beta_1' \Sigma_{10}) \Sigma_{00}^{-1} \alpha_1 + \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{01} \beta_1) - \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{00}) \Sigma_{00}^{-1} \alpha_1$$

Applying (C.6) to the latter expression, this becomes

$$= 2\lambda_1 \partial(\beta_1' \Sigma_{11} \beta_1) + (\beta_1' S_{1\epsilon} \Sigma_{00}^{-1} \alpha_1 + \alpha_1' \Sigma_{00}^{-1} S_{\epsilon 1} \beta_1) - \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{00}) \Sigma_{00}^{-1} \alpha_1$$

such that from (C.5) we get

$$\partial \lambda_1 = \lambda_1 \partial(\beta_1' \Sigma_{11} \beta_1) + (\beta_1' S_{1\epsilon} \Sigma_{00}^{-1} \alpha_1 + \alpha_1' \Sigma_{00}^{-1} S_{\epsilon 1} \beta_1) - \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{00}) \Sigma_{00}^{-1} \alpha_1 \quad (C.9)$$

Now eliminating Σ_{11} and $S_{\epsilon 1}$ from (C.8) and (C.9) we find that

$$\partial \lambda_1 = -\lambda_1^{-1} \alpha_1' \Sigma_{00}^{-1} (\partial \Omega) \Sigma_{00}^{-1} \alpha_1 - (1 - \lambda_1^{-1}) \alpha_1' \Sigma_{00}^{-1} (\partial \Sigma_{00}) \Sigma_{00}^{-1} \alpha_1 \quad (C.10)$$

which completes the proof of Theorem 1. \square

References

- Ahn, S.K. & Reinsel, G.C. (1990): "Estimation for partially non-stationary multivariate autoregressive models". *Journal of the American Statistical Association*, 85, 813-823.
- Anderson, T.W. (1951): "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions". *Annals of Mathematical Statistics*, 22, 327-351.
- Anderson, T.W. (1971): *The Statistical Analysis of Time Series*. New York Wiley.
- Andrews, D.W.K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica*, 59, 817-858.
- Bartlett, M.S. (1946): "On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series. *Journal of the Royal Statistical Society, Supplement* 8, 27-41, 85-97, and (1948) 10, 200.
- Engle, R.F. & Granger, C.W.J. (1987): "Co-integration and Error-correction: Representation, Estimation and Testing". *Econometrica*, 55, 251-276.
- Hansen, H. (1991): "Maksimum likelihood kointegrationsanalyse af den danske løndannelse". Institute of Economics, University of Copenhagen.
- Johansen, S. (1988): "Statistical Analysis of Cointegration Vectors". *Journal of Economic Dynamics and Control*, 12, 231-254.
- Johansen, S. (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 59, 1551-1580.
- Johansen, S. & Juselius, K. (1990): "Maximum Likelihood Estimation and Inference on Cointegration - With applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, 52, 169-210.
- Johansen, S. & Juselius, K. (1992): "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK.", *Journal of Econometrics*, Vol. 53, 211-244.
- Reinsel, G.C. & Ahn, S.K. (1990): "Vector AR Models with Unit Roots and Reduced Rank Structure: Estimation, Likelihood Ratio Tests, and Forecasting". Discussion Paper, University of Wisconsin.

Discussion Papers from Institute of Economics, University of Copenhagen

- 91-01 Flemming Christensen: On the Nucleolus of General Cooperative Games.
 - 91-02 Y. P. Gupta: Bayesian Analysis II - Applications to Danish Data.
 - 91-03 Søren Bo Nielsen: A Note on the Sustainability of Primary Budget Deficits.
 - 91-04 Karl Gunnar Persson: The Never Ending Controversy: Agrarian and Industrial Productivity Growth in Britain 1750-1860.
 - 91-05 Katarina Juselius: Domestic and Foreign Effects on Prices in an Open Economy.
 - 91-06 Lars Haagen Pedersen & Peter Stephensen: A Decentralized Ramsey Growth Model with a Monopoly Union.
 - 91-07 Flemming Christensen: A Kernel Variant for General Cooperative Games: Characterization Results.
 - 91-08 Hans Aage: The Baltic Republics.
 - 91-09 Hans Aage: Convergence of Preferences, Proposals, and Prospects.
 - 91-10 Jens Bonke: Distribution of Economic Resources - Implications of Including the Household Production.
 - 91-11 Karl Vind: Figure 4.
 - 91-12 Hans Jørgen Jacobsen & Christian Schultz: Decreasing Unemployment Increases Welfare.
 - 91-13 John Driffill & Christian Schultz: Renegotiation in Repeated Cournot-Duopoly.
 - 91-14 Katarina Juselius: On the Design of Experiments when Data Are Collected by Passive Observation.
 - 91-15 Katarina Juselius: On the Duality between Long Run Relations and Common Trends in an Empirical Analysis of Aggregate Money Holdings.
 - 91-16 Hans Jørgen Jacobsen & Christian Schultz: On the Effectiveness of Economic Policy when Competition is Imperfect and Expectations are Rational.
 - 91-17 Birgitte Sloth: Voting over the Partition of a Pie.
 - 91-18 Katarina Juselius: Long-run Relations in Australian Monetary Data.
 - 91-19 Hans Jørgen Jacobsen & Christian Schultz: On Price Discrimination by Rationing.
 - 91-20 Troels Østergaard Sørensen: Labor Supply and the Equilibrium Rate of Unemployment.
 - 91-21 Per Callesen: Comparative Properties of ADAM and the Treasury Model.
 - 91-22 Flemming Christensen & Jens Leth Hougaard: Input Measures of Technical Efficiency: An Axiomatization.
 - 91-23 Torben Tranæs: Common Agency and Decision Making - A Note on Existence.
- ***
- 92-01 Lisbeth Funding la Cour: A Divisia Money Demand Analysis for Denmark. An Application of the "Johansen-procedure" for Estimating Multivariate Cointegrated Systems.
 - 92-02 Niels Thygesen: Economic and Monetary Union: Critical Notes on the Maastricht Treaty Revisions.

- 92-03 Mun S. Ho, William R. M. Perraudin & Bent E. Sørensen: Multivariate Tests of a Continuous Time Equilibrium Arbitrage Pricing Theory with Conditional Heteroskedasticity and Jumps.
- 92-04 Søren Johansen & Katarina Juselius: Identification of the Long-Run and the Short-Run Structure. An Application to the ISLM Model.
- 92-05 Ebbe Hendon, Hans Jørgen Jacobsen, and Birgitte Sloth: The One-Deviation Principle for Sequential Rationality.
- 92-06 Christian Schultz: A Note on Strongly Renegotiation-Proof Equilibria.
- 92-07 Eskil Heinesen: CES Transaction Functions in Macroeconomic Rationing Models.
- 92-08 Eskil Heinesen: A Note on the Simple Two-Variable CES Transaction Function in Macroeconomic Rationing Models.
- 92-09 Ebbe Hendon, Hans Jørgen Jacobsen, Birgitte Sloth, and Torben Tranæs: Expected Utility under Uncertainty.
- 92-10 Ebbe Hendon, Hans Jørgen Jacobsen, Michael Teit Nielsen, and Birgitte Sloth: Learning, Tracing and Risk Dominance.
- 92-11 Karl Gunnar Persson: Regional and National Productivity Comparisons without Production Data. Method and an Application to 14th Century Europe.
- 92-12 Torben Tranæs: Path Perfectness in Games of Perfect Information.
- 92-13 Henrik Hansen & Søren Johansen: Recursive Estimation in Cointegrated VAR-Models.

ISSN 0902-6452